Stat 350: Lab 3

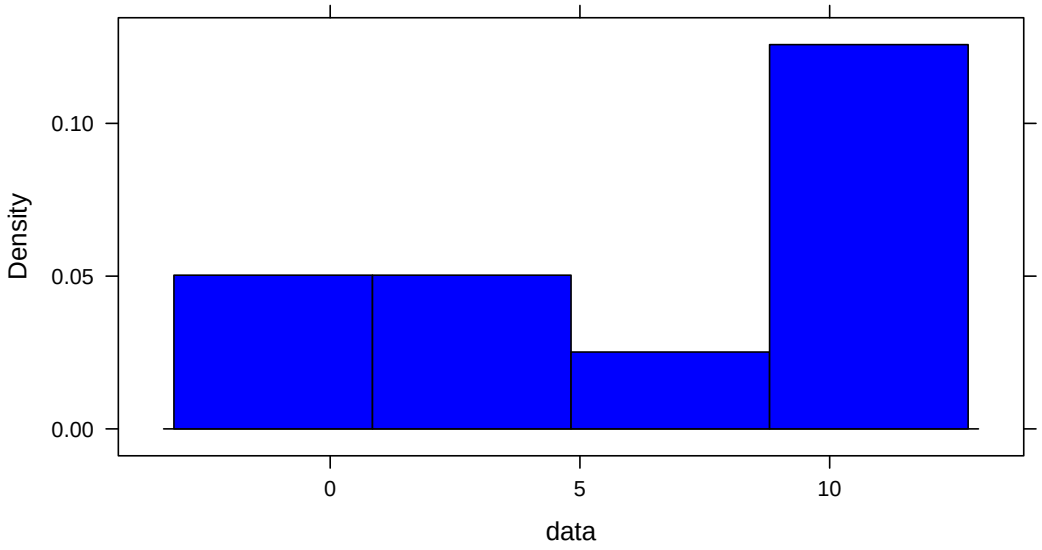
James Watterson

February 4, 2015

A Generated Sample of 10

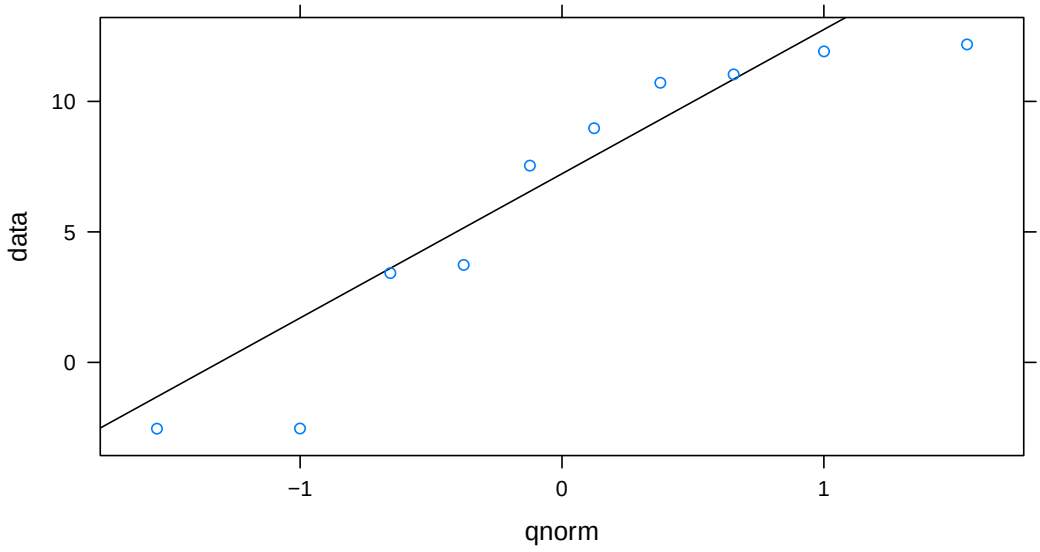
1. The histogram roughly follows the normal density curve but because there are so few units, it is a very rough approximation.

```
> my_histogram(nums_10)
```



2. The plot does not show any important deviations, the values follow roughly the linear line that goes through the first and third quantiles.

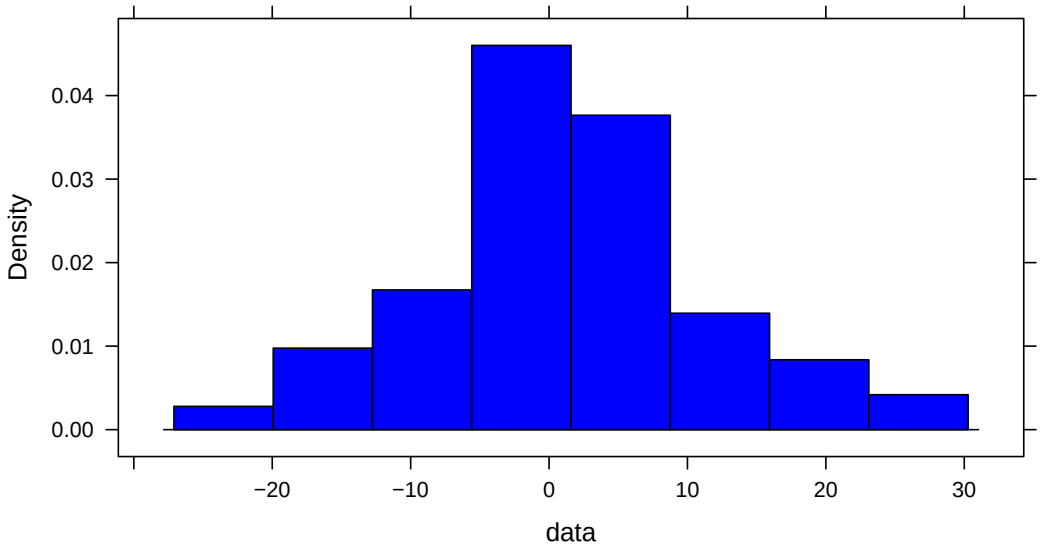
```
> my_qqwithline(nums_10);
```



B Generated Sample of 100

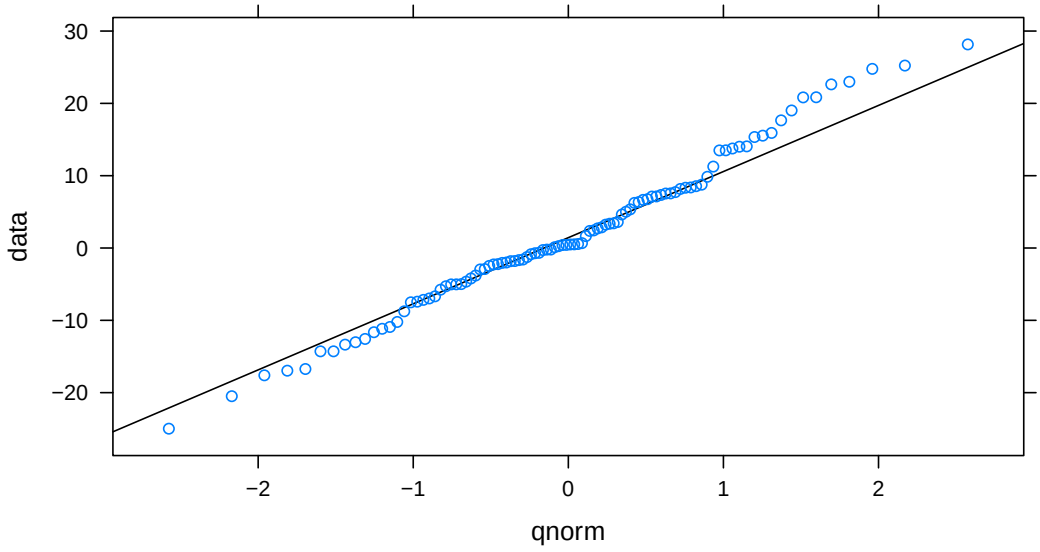
1. The histogram follows very closely the normal curve, with a distinctive hump in the middle.

```
> my_histogram(nums_100)
```



2. The plot does not show any important deviations, the values follow very closely the linear line that goes through the first and third quantiles

```
> my_qqwithline(nums_100);
```

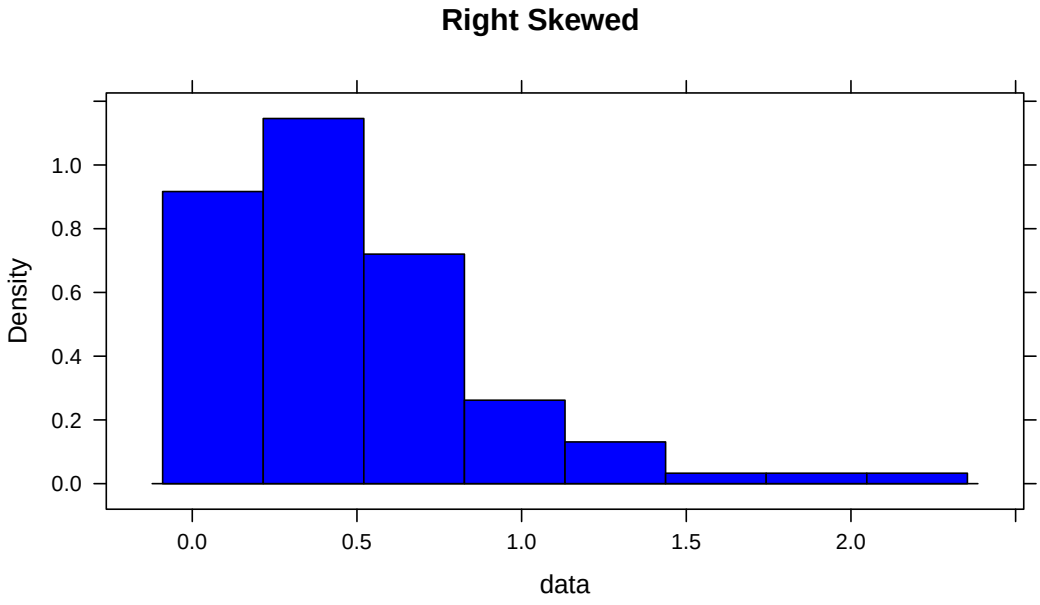


3. The plots from A and B may have different values but they are the same plot to different degrees of granularity because the values are all from the same normal distribution. In general plots from B are better because they have more units of data.

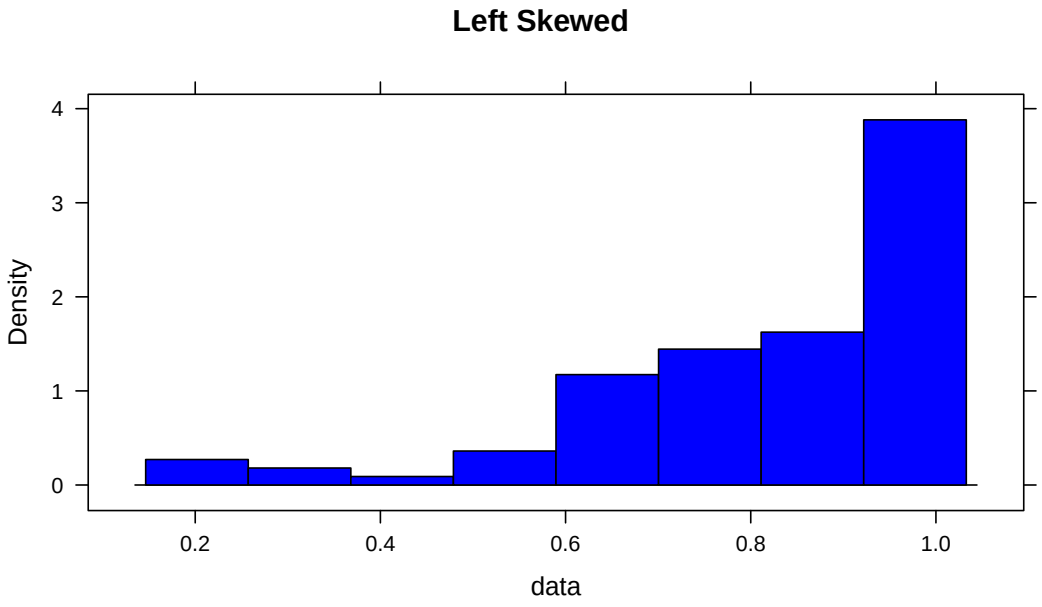
C Skews and Tails in Data

1. Skewed and tailed data produced histograms as expected, the left and right skews showed a clustering on the upper and lower end of the values respectively. The short tail data was over a shorter range and the long tail data was over a much larger range.

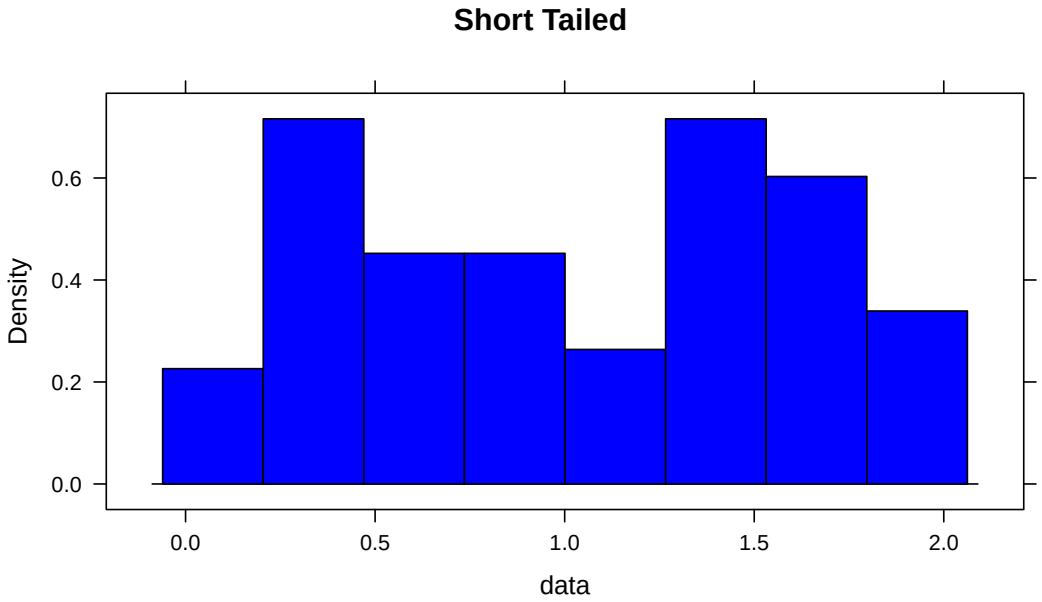
```
> my_histogram(right_skew, "Right Skewed")
```



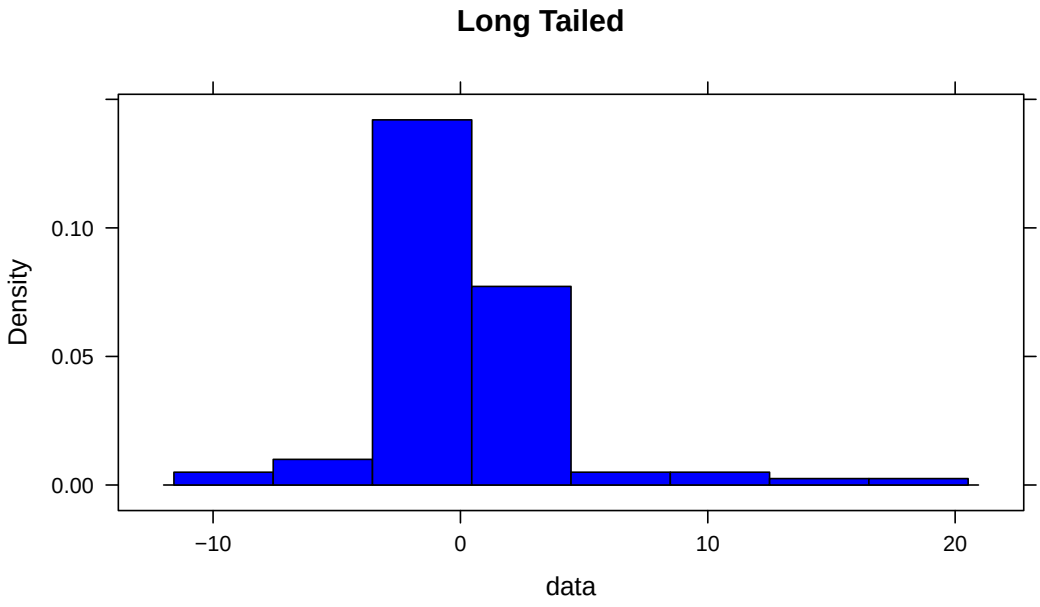

```
> my_histogram(left_skew, "Left Skewed")
```




```
> my_histogram(short_tail, "Short Tailed")
```

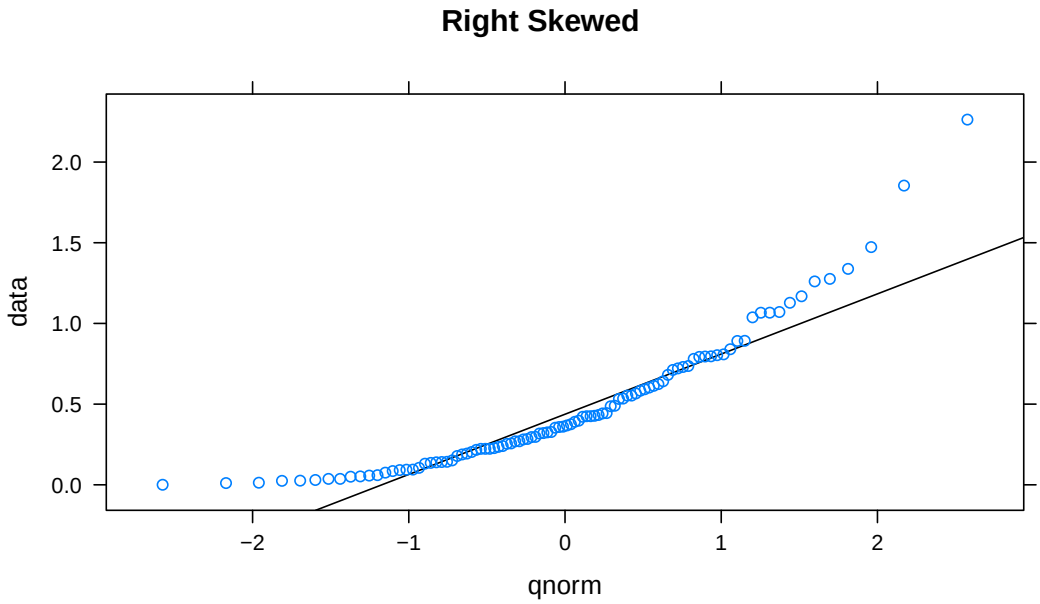



```
> my_histogram(long_tail, "Long Tailed")
```

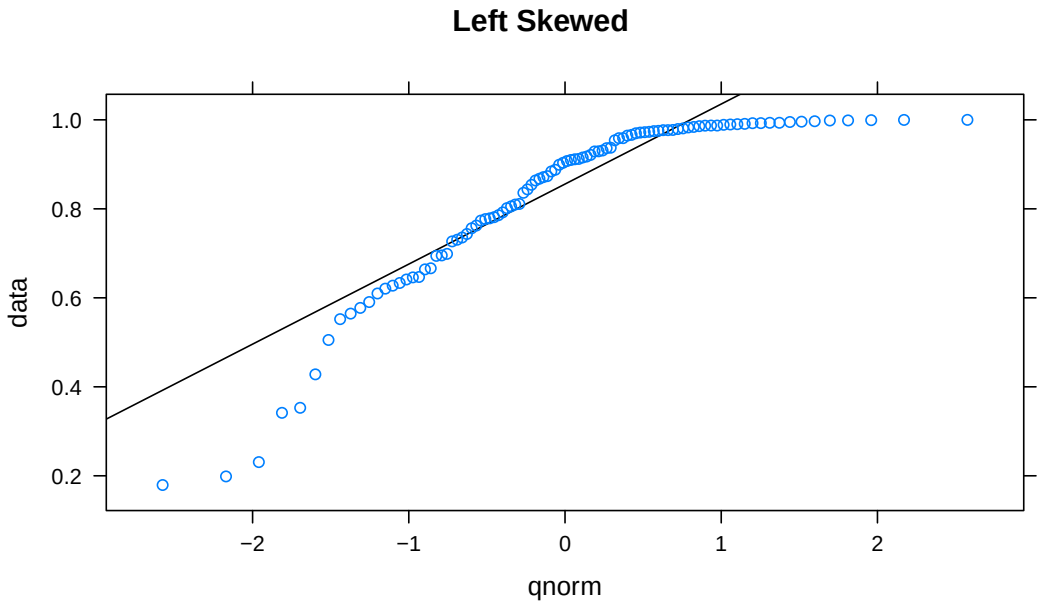


2. Skewed and tailed data produced normal quantile plots as expected, the left and right skews bent towards logarithmic and exponential curves respectively as opposed to the standard nearly linear curve of unskewed data. Similarly, the long tail data was much more flat than normal and the short tail data much more steep.

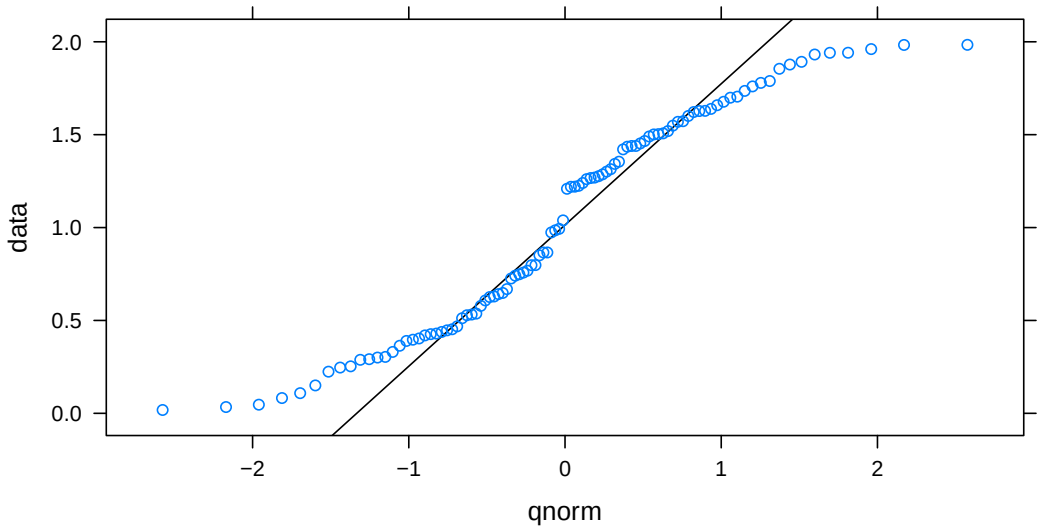
```
> my_qqwithline(right_skew, "Right Skewed")
```



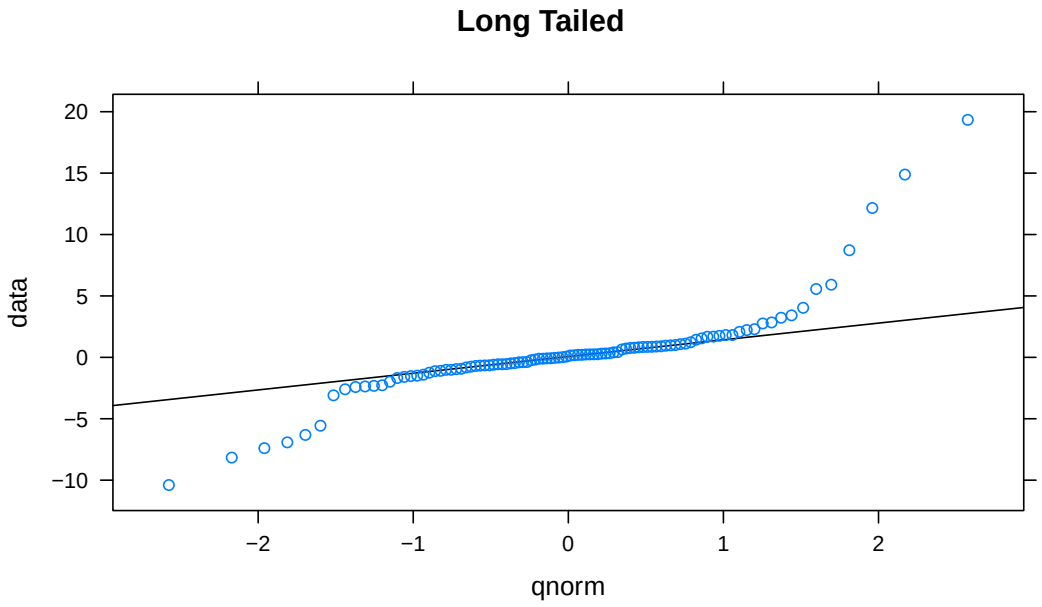

```
> my_qqwithline(left_skew, "Left Skewed")
```




```
> my_qqwithline(short_tail, "Short Tailed")
```



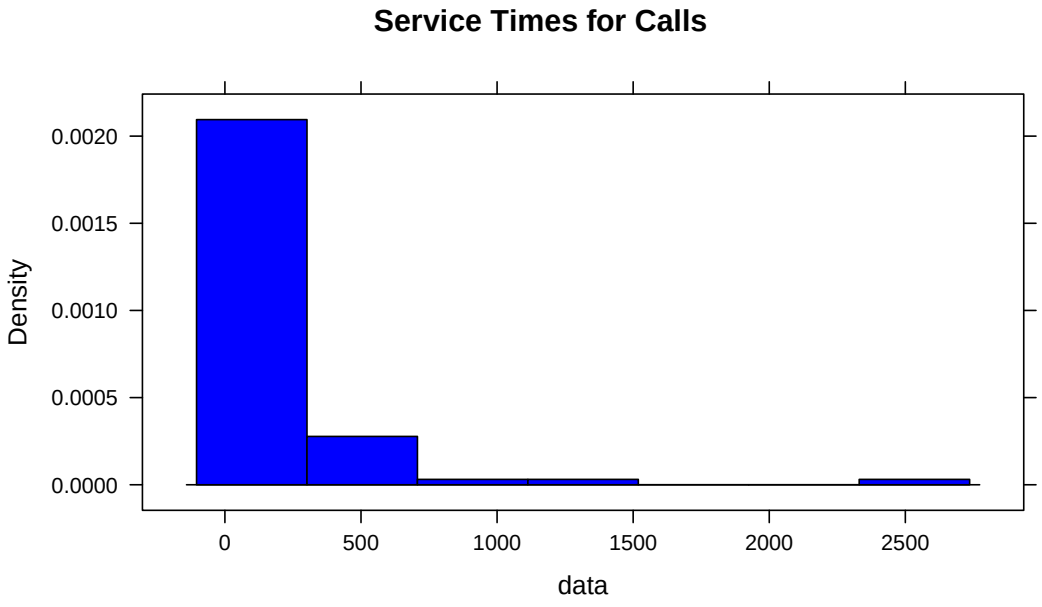

```
> my_qqwithline(long_tail, "Long Tailed")
```



D Call Center Data

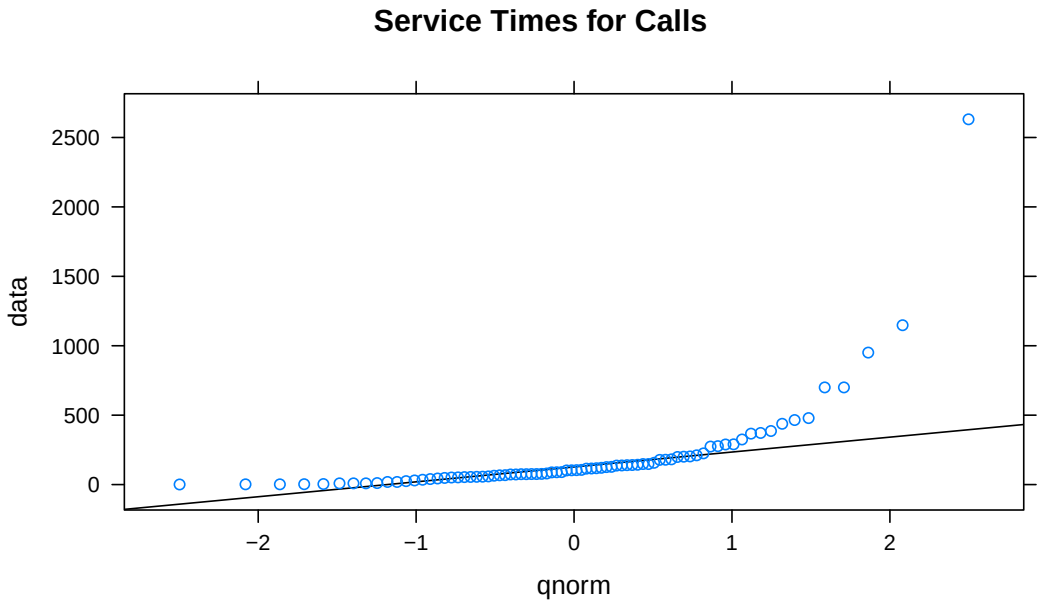
1. Based on the histogram, this data appears to be either right skewed or long tailed. The clustering of data near the low values is what implies whereas the large range is the source of my long tailed suspicions, either way the quantile plot should be able to confirm.

```
> my_histogram(call_data, "Service Times for Calls")
```



2. The quantile plot confirms the fact that this is a long tailed distribution. Although the curve bends slightly towards expenontial (a sign of a right skew) the much diminished slope (compared to a standard linear curve of y=x) such that it is almost flat, confirms suspicions that it is in fact long tailed.

```
> my_qqwithline(call_data[,1], "Service Times for Calls")
```



Appendix: Data and Functions

Functions:

```
my_qqwithline <- function(data, title=NULL){
  qqmath(data,
         panel = function(x) {
           panel.qqmathline(x, distribution = qnorm)
           panel.qqmath(x)
         },
         main=title)
}
my_histogram <- function(data, title=NULL){
  histogram(~data, type="density", col="blue", main=title)
}
```

Generation of Data:

```
nums_10 <- rnorm(10, mean=2, sd=10)
nums_100 <- rnorm(100, mean=2, sd=10)
n <- 100
right_skew <- rexp(n,rate=2)
left_skew <- rbeta(n,2,0.5,ncp=2)
short_tail <- runif(n,min=0,max=2)
long_tail <- rcauchy(n,location=0,scale=1)
```

Data:

	1	2	3	4	5	6	7	8	9	10
1	11.92	-2.53	7.54	10.71	3.42	-2.54	12.18	3.73	8.97	11.04

Table 1: nums_10

	1	2	3	4	5	6	7	8	9	10
1	20.85	-5.28	0.49	22.63	8.38	5.02	6.75	-4.65	-5.76	-0.28
2	8.76	-2.47	13.50	-0.85	-5.03	-20.49	-6.97	4.64	3.23	0.51
3	15.35	7.34	-0.21	1.65	-0.72	-1.67	-4.20	-7.17	2.37	-4.98
4	0.42	17.65	6.36	20.83	-2.00	22.99	-1.81	-2.96	-7.42	-10.92
5	2.72	-24.98	19.02	3.42	7.54	11.26	14.01	-5.02	-1.80	0.56
6	-2.23	7.14	-14.28	-1.59	8.17	2.44	-1.25	-13.03	3.60	-17.60
7	24.78	-3.80	-10.22	-11.17	0.43	-2.27	-0.66	7.12	-2.91	-11.64
8	6.66	2.86	0.68	-13.36	-16.73	8.34	15.92	14.06	0.26	0.12
9	7.73	25.23	-16.96	-8.75	13.76	28.15	7.55	-14.28	13.51	5.35
10	15.54	-0.21	8.55	3.36	-12.57	-6.68	6.24	-2.06	9.85	-7.51

Table 2: nums_100

	1	2	3	4	5	6	7	8	9	10
1	0.05	0.79	0.84	0.43	0.05	0.15	0.09	0.33	0.78	0.02
2	0.80	0.26	0.30	2.26	0.33	0.57	0.49	1.26	0.09	0.72
3	0.01	0.06	0.81	0.38	1.28	0.24	0.35	0.22	0.49	0.14
4	0.43	0.09	0.43	0.89	0.40	0.22	0.24	0.58	1.04	0.27
5	0.20	0.27	0.32	0.11	0.18	1.13	0.64	1.07	0.06	0.60
6	0.00	0.44	0.23	0.03	0.73	1.07	0.89	0.19	0.14	0.25
7	0.22	0.08	0.61	0.13	0.53	0.55	1.34	0.32	0.01	0.55
8	0.30	0.71	0.62	1.85	0.44	0.04	0.59	0.74	0.03	0.43
9	0.42	0.80	0.36	0.28	0.28	0.36	0.53	0.37	1.07	0.22
10	0.19	1.17	0.39	0.14	0.80	0.68	0.14	0.04	1.47	0.09

Table 3: right_skew

	1	2	3	4	5	6	7	8	9	10
1	0.97	0.65	0.91	0.91	0.97	0.69	0.56	1.00	0.35	0.99
2	0.87	0.70	0.20	0.92	1.00	0.98	0.98	0.97	0.87	0.97
3	0.93	0.63	0.63	0.84	1.00	0.76	0.99	0.89	0.65	0.97
4	0.90	0.96	0.98	0.99	0.93	0.81	0.73	0.81	0.99	0.78
5	0.92	0.96	0.55	1.00	0.93	0.98	0.74	0.78	0.23	0.99
6	0.99	1.00	0.99	0.67	0.43	0.64	0.84	0.91	0.79	0.66
7	0.81	0.98	0.98	0.96	0.79	0.76	1.00	1.00	1.00	0.77
8	0.18	0.99	0.99	0.62	0.86	0.74	0.91	0.51	0.98	0.99
9	0.94	0.61	0.80	0.59	0.98	0.34	0.94	0.70	0.97	0.90
10	0.99	0.95	0.92	0.58	0.85	0.99	0.88	0.73	0.87	0.78

Table 4: left_skew

	1	2	3	4	5	6	7	8	9	10
1	1.44	1.27	1.57	1.94	0.97	1.62	1.66	0.54	0.47	1.22
2	1.98	1.76	0.43	0.43	1.29	0.25	0.29	1.04	0.42	0.63
3	1.60	0.87	1.70	1.98	0.53	0.77	1.79	1.78	1.63	1.21
4	0.22	0.45	1.24	0.25	0.45	1.34	1.85	1.44	0.08	0.53
5	0.03	1.47	0.44	0.67	0.05	0.02	0.87	1.55	0.40	0.99
6	1.42	1.22	0.39	1.30	0.64	0.33	1.22	1.51	1.52	0.73
7	1.27	1.49	0.74	0.65	1.89	1.96	0.29	0.85	0.30	0.40
8	0.11	0.76	1.35	0.15	0.80	1.57	1.74	1.50	1.50	1.44
9	1.70	1.63	0.63	1.88	0.51	1.68	0.75	0.36	0.80	1.93
10	1.31	1.28	1.64	0.61	1.26	0.58	0.30	0.99	1.94	1.45

Table 5: short_tail

	1	2	3	4	5	6	7	8	9	10
1	2.76	-0.56	0.01	2.08	0.31	-0.39	-0.02	0.06	-0.23	-0.07
2	-8.16	1.00	2.22	-0.65	-2.41	1.23	-0.04	-0.08	-1.49	0.25
3	-0.96	-0.70	-0.76	1.55	2.29	-1.10	-7.39	-5.56	1.67	0.24
4	1.07	-1.98	-1.52	0.77	-0.11	1.82	-0.82	2.84	-2.37	-0.60
5	-3.09	-1.68	-0.55	0.85	12.16	4.03	-10.39	3.42	3.23	0.25
6	5.57	14.88	0.17	-1.01	-2.60	1.44	0.72	0.82	1.75	0.90
7	8.72	-0.18	0.88	-6.31	0.78	0.30	-2.32	-1.59	0.84	0.20
8	0.33	0.95	1.69	-1.41	0.98	0.41	-2.27	0.64	1.82	-0.56
9	-1.12	-0.38	-6.92	0.15	-0.40	-1.02	-0.47	-0.66	-0.93	1.12
10	-0.49	-1.25	-0.11	5.92	0.86	0.24	-0.67	19.34	0.20	0.42

Table 6: long_tail